AD-A042 683 FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS F/G 12/1
A DIFFERENTIAL FOR L-STATISTICS.(U)
JUN 77 D D BOOS N00014-76-C-0608
UNCLASSIFIED FSU-STATISTICS-M416 ONR-TR-116 NL

1 OF 1

ADA042 683

END
DATE
FILMED
8 - 77
DDC

12

6 A DIFFERENTIAL FOR L-STATISTICS

by


10 Dennis D. Boos


14 FSU-STATISTICS ~~REPORT~~ M416
ONR Technical Report No. 116

18 ONR 19 TR-116

9 Technical rept.,

12 41p.




11 June, 1977
Department of Statistics
Florida State University
Tallahassee, Florida 32306



15 Research supported by the Army, Navy, and Air Force under Office of Naval Research Contract No. N00014-76-C-0608. Reproduction in whole or in part is permitted for any purpose of the United States Government.




400277

# ABSTRACT

## A DIFFERENTIAL FOR L-STATISTICS

Let $X_{1n} \le X_{2n} \le \ldots \le X_{nn}$ be an ordered sample from a distribution F and $c_{in}$ a sequence of constants. Statistics of the form $\sum_{i=1}^{n} c_{in} X_{in}$ are called "linear functions of order statistics," "L-estimators," or simply "L-statistics." Various methods of generating the constants $c_{in}$ have been considered, including $\tilde{c}_{in} = n^{-1} J(i/n+1)$ or $n^{-1} J(i/n)$ and $\hat{c}_{in} = \int_{(i-1)/n}^{i/n} J(u)du$ for fixed "score" functions J. L-statistics of the form $T_n = \sum_{i=1}^{n} \hat{c}_{in} X_{in}$ can be obtained from the functional $T(F) = \int F^{-1}(t) J(t) dt$ by substitution of the sample d.f. $F_n$ for F. Under the mild assumption that J is bounded and continuous a.e. Lebesgue and a.e. $F^{-1}$ and under a tail restriction on F of the form $\int q(F(x))dx < \infty$ (e.g., $q(t) = [t(1-t)]^{\frac{1}{2}-\delta}$, $0 < \delta < \frac{1}{2}$), it is shown that $T(\cdot)$ has a Frechet-type differential. The tail restriction may be dropped if J trims the extremes. In either case it follows that if $\{X_i\}$ is a sequence of independent observations on F, then $\sqrt{n}(T(F_n) - T(F))$ is asymptotically normal and obeys a law of the iterated logarithm. Continuity of T holds under milder conditions on J and F. This leads to strong consistency, $T(F_n) \xrightarrow{wp1} T(F)$. No continuity restrictions are imposed on F, so that the results are applicable to a wide class of distributions of interest in robust estimation. Illustration is provided by examples including the trimmed mean, the smoothly trimmed mean, and approximations to the interquartile range. The asymptotic normality result is competitive with one of Stigler (1974) for the closely related statistic $S_n = \sum_{i=1}^{n} \tilde{c}_{in} X_{in}$, obtained under stronger conditions on J but a slightly milder condition on F. However, in addition to asymptotic normality of $T(F_n)$,

the differential approach of the present paper yields characterization of the almost sure behavior of $T(F_n)$ and lends itself to straightforward extension to the case of dependent variables.

## 1. INTRODUCTION

Let $X_{1n} \leq X_{2n} \leq \ldots \leq X_{nn}$ be an ordered sample from a distribution F and $c_{in}$ a sequence of constants. Statistics of the form $\sum_{i=1}^{n} c_{in} X_{in}$ are called "linear functions of order statistics" (e.g., Shorack (1972), Stigler (1974)), "L-estimators" (Huber (1972)), or simply "L-statistics". Various methods of generating the constants $c_{in}$ have been considered, including $\tilde{c}_{in} = n^{-1}J(i/n)$ or $n^{-1}J(i/n+1)$ and $\hat{c}_{in} = \int_{(i-1)/n}^{i/n} J(u)du$ for fixed "score" functions J. By choosing the $c_{in}$ (or J) properly, the L-estimator can be made insensitive to outliers or "robust" with regard to long-tailed distributions. Thus L-estimators have played an important role in the development of the modern theory of robust estimation. Since their exact sampling distributions are difficult to compute even under strict model assumptions, much work has focused upon the question of asymptotic normality. There appear to be four main approaches: (1) Weak convergence in connection with the empirical process, Bickel (1967), Shorack (1969, 1972); (2) Hajek's projection method, Stigler (1969, 1974); (3) approximation of the L-estimator by sums of independent exponential r.v.'s, Chernoff, Gastwirth, and Johns (1967); (4) Frechet differentials, Gregory (1976).

The present paper treats L-statistics of the form $T_n = \sum_{i=1}^{n} \hat{c}_{in} X_{in}$ and uses a differential method similar to Gregory's, but the restrictions on the $\hat{c}_{in}$ and the underlying d.f. F are similar in spirit to Stigler (1974) who treats the closely related statistic $S_n = \sum_{i=1}^{n} \tilde{c}_{in} X_{in}$. Further, in addition to asymptotic normality, the main theorems of Sections 4 and 5 yield strong consistency and a law of the iterated logarithm (LIL). More general versions of these last two results have been given recently by Wellner (1977a), (1977b).

The L-statistics we consider may be represented in terms of a functional defined on the space of d.f.'s. The basic functional of interest is given by

(1.1) $$T(F) = \int_0^1 F^{-1}(t)dK(t) ,$$

where $F^{-1}(t) = \inf\{x: F(x) \geq t\}$ and $K(t)$ is a right continuous function of bounded variation on $[0,1]$. In this paper we restrict $K(t)$ to be absolutely continuous by putting $K(t) = \int_0^t J(u)du$, for $J$ integrable on $[0,1]$. Then (1.1) becomes

(1.2) $$T(F) = \int_0^1 F^{-1}(t)J(t)dt .$$

(Note that finite linear combinations of quantiles are thus excluded. However, in Example (iii) of Section 4 we show how to approximate such combinations by appropriate choices of $J$.) For the case that $K(t)$ is a d.f. symmetric about $\frac{1}{2}$, Bickel and Lehmann (1976, I and II) discuss the value of (1.1) as a measure of location. We are allowing $K(t)$ to represent a signed measure in order to include a large class of scale functionals as well.

The sample d.f. $F_n$ generated by a sample $X_1, \ldots, X_n$ is defined by

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x) .$$

For estimation purposes we set $T_n = T(F_n)$ by substitution of $F_n$ for $F$ in (1.1). This statistic can be put in a more familiar form by noting that

$$\int_0^1 F^{-1}(t)dK(t) = \int_{-\infty}^{\infty} xdK(F(x))$$

for any d.f. in the domain of $T(\cdot)$ (the change of variable is justified by Lemma 12, Section 3). Then substitution of $F_n$ for $F$ in this last expression yields

(1.3) $$T(F_n) = \int_{-\infty}^{\infty} xdK(F_n(x)) = \int_{-\infty}^{\infty} x \, d\left[\int_0^{F_n(x)} J(u)du\right] = \sum_{i=1}^{n} X_{in} \int_{\frac{i-1}{n}}^{\frac{i}{n}} J(u)du,$$

i.e., $T(F_n) = T_n$.

In Sections 2 and 3 some preliminaries are developed. Section 2 provides the definitions of the differential and of continuity, the basic tools to be utilized in the sequel, and important convergence results for $F_n - F$. Of interest here are recent results by O'Reilly (1974), James (1975), and Wellner (1977a) for weighted empirical processes. Statistical application is provided by Theorem 1 (strong consistency), Theorem 2 (asymptotic normality), and Theorem 3 (LIL). Section 3 provides a useful representation for the difference $T(F) - T(G)$ and a related inequality.

In Sections 4 and 5 we will prove that under suitable restrictions on J, F, and the sequence $\{X_i\}$, we have

(1.4) $$\lim_{n\to\infty} T(F_n) = T(F) \quad \text{w.p.1} \;;$$

(1.5) $$\sqrt{n}(T(F_n) - T(F)) \xrightarrow{L} N(0, \sigma^2) \text{ as } n \to \infty \;;$$

(1.6) $$\overline{\lim_{n\to\infty}} \frac{\sqrt{n}(T(F_n) - T(F))}{\sqrt{2\sigma^2 \log\log n}} = 1 \quad \text{w.p.1.}$$

Continuity of T leads to (1.4), and the existence of a differential for T leads to (1.5) and (1.6). For proving continuity we assume that J is bounded and integrable on [0,1]. For proving the existence of a differential we assume that J is bounded and continuous a.e. Lebesque and a.e. $F^{-1}$. In Section 4 we further restrict J to be 0 in some neighborhoods of 0 and 1, making T *robust* in the Hampel (1971) sense. Theorem 4 provides the existence of a differential in this case w.r.t. the sup-norm $||\cdot||_\infty$. A corollary yields (1.5) and (1.6) for I.I.D. r.v.'s by appeal to Theorems 2 and 3. Theorem 5 yields continuity of T w.r.t. $||\cdot||_\infty$ and (1.4). In Section 5 we trade the trimming condition on J for a tail restriction on F of the form

$\int q(F(x))dx < \infty$ (e.g., $q(t) = (t(1-t))^{\frac{1}{2}-\delta}$, $0 < \delta < \frac{1}{2}$). Theorem 6 then provides the existence of a differential w.r.t. the q-norm $||\cdot||_{q(F)}$, essentially defined by $||\cdot||_{q(F)} = ||(\cdot)/q(F)||_\infty$. Corollaries 1 and 2 then yield (1.5) and (1.6) in the I.I.D. case for different classes of q functions (tail restrictions). Continuity and (1.4) are provided by Theorem 7.

Section 6 generalizes the results of Sections 4 and 5 to the functional $T(F) = \int_0^1 h(F^{-1}(t))dK(t)$. In Section 7 specific comparisons are made with related work of Stigler (1974), Gregory (1976), and Wellner (1977a), (1977b).

## 2. THE DIFFERENTIAL AND ITS STATISTICAL APPLICATIONS

Most of this section can be found in expanded form in Boos and Serfling (1977). We begin with the definition of the differential and motivation for its use. Lemmas 1-4 spell out the statistical properties of the differential and Lemma 5 relates continuity of functionals to consistency of $T(F_n)$. A discussion of specific norms follows, including Lemmas 6-11 which provide some convergence properties of $||F_n - F||$ for these norms. Lastly, Theorems 1, 2, and 3 provide (1.4), (1.5), and (1.6) for the usual situation of I.I.D. r.v.'s.

Let T be a real-valued functional defined on a *convex* set $F$ of d.f.'s. Denote by $\mathcal{D}(F)$ the linear space generated by *differences* H-G of members of $F$, i.e.,

$$\mathcal{D}(F) = \{\Delta:\ \Delta = a(H-G),\ H,G \in F,\ a \in R\}.$$

Assume $\mathcal{D}(F)$ is equipped with a norm $||\cdot||$, to be specified later.

DEFINITION. We say that a functional T defined on $F$ has a *differential* at the point $F \in F$ w.r.t. the norm $||\cdot||$ and the set $G_F \in F$ if there exists a quantity $T(F; \Delta)$ defined on $\Delta \in \mathcal{D}(F)$, which is linear in the argument $\Delta$ and satisfies the condition

$$\lim_{\substack{||G-F|| \to 0 \\ G \in \mathcal{G}_F}} \frac{T(G) - T(F) - T(F; G-F)}{||G-F||} = 0 \,. \; \square \tag{2.1}$$

$T(F; \Delta)$ is called the "differential." By *linearity* of $T(F; \Delta)$ is meant that

$$T(F; \sum_{i=1}^{k} a_i \Delta_i) = \sum_{i=1}^{k} a_i T(F; \Delta_i)$$

for $\Delta_1, \ldots, \Delta_k \in \mathcal{D}(F)$ and real $a_1, \ldots, a_k$. Often $\mathcal{G}_F = F$, although sometimes (2.1) is easier to show for special choices of $\mathcal{G}_F$. When mention of $\mathcal{G}_F$ is omitted, it will be assumed that $\mathcal{G}_F = F$.

The intuitive content of the above definition is that $T(G) - T(F)$ can be closely approximated by the differential $T(F; G-F)$, whose linearity property can then be exploited. An alternative statement of (2.1) is

$$T(G) - T(F) = T(F; G-F) + o(||G-F||) \text{ as } ||G-F|| \to 0,\ G \in \mathcal{G}_F. \tag{2.2}$$

For statistical applications we substitute the sample d.f. $F_n$ for $G$ and find that $T(F_n) - T(F)$ is approximated closely by $T(F; F_n - F)$ in a stochastic sense (to be made clear in the lemmas below). In order to examine $T(F; F_n-F)$, recall that for a sample $X_1, \ldots, X_n$, the sample d.f. may be written as $F_n = n^{-1} \sum_{i=1}^{n} \delta_{X_i}$, where $\delta_x$ denotes the d.f. of a r.v. degenerate at $x$. Then, by the linearity of the differential,

$$T(F; F_n-F) = T(F; \frac{1}{n} \sum_{i=1}^{n} (\delta_{X_i} - F)) = \frac{1}{n} \sum_{i=1}^{n} T(F; \delta_{X_i} - F) \,.$$

We assume throughout this paper that $T(F; \delta_x - F)$ is well-defined and measurable w.r.t. the probability space induced by X. For convenience we set $T(F;\ \delta_x\text{-}F) = T[F;\ x]$ and note that $T[F;\ x] = IC_{T,F}(x)$, the influence curve of T at F (see Hampel (1974)). (In Boos and Serfling (1977), $T[F;\ x] = T(F;\ \delta_x\text{-}F) - \int T(F;\ \delta_x\text{-}F)dF(x)$; this latter expectation is 0 for L-statistics.) Thus we see that $T(F;\ F_n\text{-}F)$ reduces to an average of identically distributed r.v.'s, $T[F;\ X_i]$, which are independent if the original sample $X_1, \ldots, X_n$ consists of independent r.v.'s.

Before pursuing the statistical applications of the differential, we formulate the related concept of continuity of functionals.

DEFINITION. The functional T is said to be *continuous* at F w.r.t. $||\cdot||$ and $\mathcal{G}_F$ if

(2.3) $$\lim_{\substack{||G-F|| \to 0 \\ G \in \mathcal{G}_F}} T(G) = T(F) .$$

In order to facilitate the statement of a number of short lemmas, we list here a group of conditions. The sample d.f. $F_n$ will always be assumed to be generated by a sample (not necessarily independent) $X_1, \ldots, X_n$ from a distribution F.

CONDITIONS.

(2.4) T has a differential at F w.r.t. $||\cdot||$ and $\mathcal{G}_F$;

(2.5) $$P\{F_n \in \mathcal{G}_F,\ \text{all } n \text{ sufficiently large}\} = 1;$$

(2.6) $$\lim_{n\to\infty} ||F_n-F|| = 0 \quad \text{w.p.1} ;$$

(2.7) $$\sqrt{n}||F_n-F|| = 0_p(1) \text{ as } n \to \infty ;$$

(2.8) $$\frac{\sqrt{n}\,||F_n-F||}{\sqrt{\log\log n}} = O(1) \text{ as } n \to \infty \text{ w.p.1 ;}$$

(2.9) $$E_F\{T[F;X]\} = 0,\ \mathrm{Var}_F\{T[F;X]\} = \sigma^2 > 0;$$

(2.10) $X_1, \ldots, X_n$ are independent and identically distributed with d.f. F.

REMARKS. (i) For some applications, conditions (2.5) and (2.6) may be replaced by weaker versions using convergence *in probability*. (ii) Condition (2.7) is equivalent to the condition that the sequence of distributions corresponding to $\sqrt{n}\,||F_n-F||$ is *tight*. (iii) Conditions (2.9) and (2.10) imply the classical central limit theorem

(2.11) $$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} T[F;X_i] \xrightarrow{L} N(0,\sigma^2) \text{ as } n \to \infty ,$$

and the classical law of the iterated logarithm

(2.12) $$\overline{\lim_{n\to\infty}} \frac{\sum_{i=1}^{n} T[F;X_i]}{\sqrt{2\sigma^2 n \log\log n}} = 1 \quad \text{w.p.1.} \ \square$$

The following lemmas provide the foundation for using the differential and the related concept of continuity in statistics. The proofs are trivial and will be ommitted for all but Lemma 4.

LEMMA 1. *If* (2.4), (2.5), *and* (2.6) *hold, then*

$$\lim_{n\to\infty} \frac{T(F_n) - T(F) - T(F;F_n-F)}{||F_n-F||} = 0 \quad \textit{w.p.1.} ,$$

*or equivalently*

$$T(F_n) - T(F) = \frac{1}{n} \sum_{i=1}^{n} T[F;X_i] + o(||F_n-F||) \textit{ as } n \to \infty \quad \textit{w.p.1.}$$

LEMMA 2. *If* (2.4), (2.5), *and* (2.7) *hold, then*

$$\sqrt{n}(T(F_n) - T(F) - T(F;F_n-F)) \xrightarrow{p} 0 \text{ as } n \to \infty .$$

LEMMA 3. *If* (2.4), (2.5), (2.7), (2.9), *and* (2.10) *hold, then*

$$\sqrt{n}(T(F_n) - T(F)) \xrightarrow{L} N(0,\sigma^2) \text{ as } n \to \infty .$$

(Note that conditions (2.9) and (2.10) are used here only to get (2.11).)

LEMMA 4. *If* (2.4), (2.5), (2.8), (2.9), *and* (2.10) *hold, then*

$$\overline{\lim_{n\to\infty}} \frac{\sqrt{n}(T(F_n) - T(F))}{\sqrt{2\sigma^2 \log\log n}} = 1 \quad w.p.1 .$$

(Note that condition (2.10) is needed here only to get (2.12).)

PROOF. Write $T(F_n) - T(F)$ as

$$T(F_n) - T(F) = [T(F_n) - T(F) - T(F;F_n-F)] + T(F;F_n-F).$$

By (2.12),

$$\overline{\lim_{n\to\infty}} \frac{\sqrt{n}T(F;F_n-F)}{\sqrt{2\sigma^2 \log\log n}} = 1 \quad \text{w.p.1.}$$

Thus it suffices to show that

$$\lim_{n\to\infty} \frac{\sqrt{n}[T(F_n) - T(F) - T(F;F_n-F)]}{\sqrt{2\sigma^2 \log\log n}} = 0 \quad \text{w.p.1.}$$

Write this last expression as

$$\frac{T(F_n) - T(F) - T(F;F_n-F)}{||F_n-F||} \cdot \frac{\sqrt{n}||F_n-F||}{\sqrt{2\sigma^2 \log\log n}} .$$

The first term converges to 0 w.p.1 by Lemma 1, and the second term is bounded w.p.1 by (2.8). □

LEMMA 5. *If* (2.3), (2.5), *and* (2.6) *hold, then*

$$\lim_{n\to\infty} T(F_n) = T(F) \quad \textit{w.p.1} .$$

The general statistical application of the differential and of continuity is well characterized by the preceding five lemmas. It should be clear how to extend Lemmas 3 and 4 to the case of dependent variables. After presenting a class of norms which satisfy (2.6) - (2.8), we will apply the above general results to get Theorems 1, 2, and 3 which provide strong consistency, asymptotic normality, and an LIL for $T(F_n)$ in the case of I.I.D. r.v.'s.

The statistical value of the theory developed in this section depends heavily on the choice of norm $||\cdot||$. In fact, the norm must serve two somewhat conflicting purposes. For satisfaction of (2.1) of the definition of the differential, we would like a relatively "large" norm, whereas conditions (2.6)-(2.8) are most easily satisfied for "small" norms. We now introduce a class of norms for which a number of useful stochastic results are available.

Let F be a fixed d.f. and let the closure of $(x_1, x_2)$ be the smallest interval (possibly infinite) containing the support $S_F$ of F. Let $\mathcal{F}$ = {d.f. G: $S_G \subset S_F$}, and let $\mathcal{D}(F)$ be the companion linear space of differences. For a bounded positive function q on (0,1), we define

$$||\Delta||_{q(F)} = \sup_{x_1<x<x_2} \left|\frac{\Delta(x)}{q(F(x))}\right|, \quad \Delta \in \mathcal{D}(F). \tag{2.13}$$

For $q(t)\equiv 1$ we get the usual sup-norm, $||\Delta||_\infty = \sup_{-\infty<x<\infty}|\Delta(x)|$, since on $\mathcal{D}(F)$ we have $\sup_{x_1<x<x_2}|\Delta(x)| = \sup_{-\infty<x<\infty}|\Delta(x)|$.

The important choices of q function are those for which $q(t) \to 0$ as $t \to 0$ and 1. They produce nonequivalent norms which are "larger" than $||\cdot||_\infty$, for

$$||\Delta||_\infty \leq ||q||_\infty ||\Delta||_{q(F)}, \quad \Delta \in \mathcal{D}(F). \tag{2.14}$$

The potential use of such norms in verifying (2.1) can be seen from the following inequality. Let $S_F = (-\infty,\infty)$ and $\int_{-\infty}^{\infty} q(F(x)dx < \infty$. Then

$$\left| \int_{-\infty}^{\infty} \Delta(x)dx \right| = \left| \int_{-\infty}^{\infty} \left(\frac{\Delta(x)}{q(F(x))}\right) q(F(x))dx \right|$$

$$\leq ||\Delta||_{q(F)} \int_{-\infty}^{\infty} q(F(x))dx .$$

A similar inequality will be employed in Section 5 to show that L-functionals have a differential w.r.t. $||\cdot||_{q(F)}$.

Even though $||\cdot||_\infty$ can be viewed as a member of the larger class of q-norms, it is advantageous to consider $||\cdot||_\infty$ by itself. Historically, results related to (2.6)-(2.8) for $||\cdot||_\infty$ generally preceded the results for $||\cdot||_{q(F)}$. For sequences of independent and identically distributed r.v.'s, condition (2.6) with $||\cdot|| = ||\cdot||_\infty$ is just the Glivenko-Cantelli Theorem. The conclusion of the following lemma implies (2.7) for $||\cdot|| = ||\cdot||_\infty$.

LEMMA 6. *Let* $\{X_i\}$ *be a sequence of independent observations on a non-degenerate distribution* F. *Then*

$$\sqrt{n}||F_n - F||_\infty \xrightarrow{L} Z_F \textit{ as } n \to \infty , \tag{2.15}$$

*where* $Z_F$ *is positive w.p.1* .

The proof of (2.15) for the case of F continuous was first given by Kolmogorov (1933). The distribution of $Z_F$ was given explicitly and seen not to depend on F. Extension to the case of F having finitely many discontinuities and not being purely atomic was obtained by Schmid (1958). Here the distribution of $Z_F$ was given explicitly; it depends upon F in the case of discontinuities. The general case is treated in Billingsley (1968), Section 16. See Boos and Serfling (1977) for further discussion.

The conclusion of the next lemma yields (2.8) for $||\cdot|| = ||\cdot||_\infty$.

LEMMA 7. *Let* $\{X_i\}$ *be a sequence of independent observations on a distribution* F. *Then*

$$\overline{\lim_{n\to\infty}} \frac{\sqrt{n}\,||F_n - F||_\infty}{\sqrt{\log \log n}} = \sup_{-\infty<x<\infty} \sqrt{2F(x)(1-F(x))} \quad w.p.1 \ . \tag{2.16}$$

The proof of (2.16) in the case of F continuous was given by Chung (1949). Extension to the case of F having discontinuities is due to Richter (1974).

Now we turn to the norm $||\cdot||_{q(F)}$. Most of the results in the literature concerning $||F_n - F||_{q(F)}$ are for F continuous. The following lemma shows how to use such results for the purpose of verifying (2.6)-(2.8) for *arbitrary* F. Let H denote the d.f. of uniform (0,1) r.v.'s.

LEMMA 8. *Let* $\{X_i\}$ *be a sequence of independent r.v.'s on a given probability space, having distribution* F. *Then a sequence* $\{U_i\}$ *of independent uniform* (0,1) *r.v.'s, with sample d.f.* $H_n$, *can be constructed such that*

$$||F_n - F||_{q(F)} \le ||H_n - H||_{q(H)} \quad w.p.1 \ , \ all \ n \ . \tag{2.17}$$

PROOF. For the given sequence $\{X_i\}$, defined on an arbitrary probability space, it is possible by means of randomization to construct uniform (0,1) r.v.'s $\{U_i\}$ such that

$$X_i = F^{-1}(U_i) \quad w.p.1 \ .$$

This is well-known, but a construction is provided in Boos (1977). Let $H_n$ be the sample d.f. of the constructed sequence $\{U_i\}$ and $F_n$ the sample d.f. of the sequence $\{X_i\}$. Then

$$F_n(x) = H_n(F(x)) \quad \text{w.p.1} ,$$

and

$$\sup_{x_1<x<x_2} \left|\frac{F_n(x) - F(x)}{q(F(x))}\right| = \sup_{x_1<x<x_2} \left|\frac{H_n(F(x)) - F(x)}{q(F(x))}\right|$$

$$\leq \sup_{0<t<1} \left|\frac{H_n(t) - H(t)}{q(t)}\right| \quad \text{w.p.1} ,$$

with equality if F is continuous. □

The bound (2.17) aids us in the following way. Suppose that condition (2.6), (2.7), or (2.8) holds for sequences of independent uniform (0,1) r.v.'s. Then for arbitrary sequences $\{X_i\}$, the conclusion of the lemma allows us to bound the quantity of interest, say $||F_n-F||_{q(F)}$ or $\sqrt{n}||F_n-F||_{q(F)}$, etc., by a quantity, $||H_n-H||_{q(H)}$ or $\sqrt{n}||H_n-H||_{q(H)}$, which satisfies the condition in question.

We are now ready to use recent results relating to (2.6)-(2.8). Lemmas 9, 10, and 11 are taken from Wellner (1977a), O'Reilly (1974), and James (1975) respectively. We attempt to preserve each author's notation. Each lemma is followed by a corollary which provides the proper extension for our application.

Let $\mathcal{H}(\uparrow)$ denote the set of all nonnegative, nondecreasing, continuous functions on [0,1] for which $\int_0^1 (1/h(t))dt < \infty$. Let $\mathcal{H}$ denote the set of all h such that $h(t) = h(1-t) = \bar{h}(t)$ for $0 \le t \le \frac{1}{2}$ and some $\bar{h}$ in $\mathcal{H}(\uparrow)$.

LEMMA 9. (Wellner). *Let* $\{U_i\}$ *be a sequence of independent uniform* (0,1) *r.v.'s. Let* $h \in \mathcal{H}$. *Then*

$$(2.18) \qquad \lim_{n\to\infty} \sup_{0\le t\le 1} \left|\frac{F_n(t)-t}{h(t)}\right| = 0 \quad w.p.1 .$$

It is apparent that the conclusion of Lemma 9 can be extended to $Q_1 = \{q:\ q$ is bounded on $[0,1]$ and $q(t) \ge h(t)\ \forall t \in [0,1]$, some $h \in \mathcal{H}\}$. The following corollary follows immediately from Lemmas 8 and 9.

COROLLARY. *Let* $\{X_i\}$ *be a sequence of independent r.v.'s having distribution* F *(not necessarily continuous). Let* $q \in Q_1$. *Then*

$$(2.19) \qquad \lim_{n\to\infty} ||F_n-F||_{q(F)} = 0 \quad w.p.1 .$$

Characterization of condition (2.7) for $||\cdot|| = ||\cdot||_{q(F)}$ is essentially provided by the following weak convergence result, found as Theorem 2 in O'Reilly (1974). Let D[0,1] be the space of real-valued functions on [0,1] having only jump discontinuities. Let $\rho_q(x,y) = \sup_{0\le t\le 1} |(x(t)-y(t))/q(t)|$. Since $(D,\rho_q)$ is not separable, O'Reilly uses "weak convergence" in the sense of Definition 2.1 of Pyke and Shorack (1968). Let $W_o(t)$ be a "tied-down" Wiener process (or "Brownian bridge").

LEMMA 10. (O'Reilly). *Let* q *be a continuous, nonnegative function on* [0,1], *bounded away from zero on* $[\alpha,1-\alpha]$ *for some* $0<\alpha<\frac{1}{2}$, *nondecreasing (non-increasing) on* $[0,\alpha]$ $([1-\alpha,1])$. *Let* $\{U_i\}$ *be a sequence of independent uniform* [0,1] *r.v.'s, and define* $U_n(t) = \sqrt{n}(F_n(t)-t)$. *Then*

$$(2.20) \qquad \int_0^1 \frac{e^{-\epsilon h_i^2}}{t}\, dt < \infty, \ \textit{for all } \epsilon > 0,\ i=1,\,2$$

*is both a necessary and sufficient condition for the weak convergence of* $U_n$ *to* $W_o$ *in* $(D,\rho_q)$ *where* $h_1(t) = t^{-\frac{1}{2}}q(t)$ *and* $h_2(t) = t^{-\frac{1}{2}}q(1-t)$.

Let $Q_2$ be the set of q functions satisfying the conditions of Lemma 10 including (2.20). Let $Q_3 = \{q: \ q$ is bounded on $[0,1]$ and $q(t) \geq q^*(t) \ \forall t \varepsilon [0,1]$, some $q^*(t) \ \varepsilon \ Q_2\}$. The following corollary follows from the proof of O'Reilly's Theorem 2.

COROLLARY. *Let* $\{X_i\}$ *be a sequence of independent r.v.'s having distribution* F *(not necessarily continuous). Let* $q \varepsilon Q_3$. *Then*

$$\sqrt{n}\,||F_n - F||_{q(F)} = O_p(1) \text{ as } n \to \infty . \tag{2.21}$$

PROOF. Following the proof of O'Reilly's Theorem 2, let $\overline{U}_n$ and $\overline{W}_o$ be versions of $U_n$ and $W_o$ defined on a common probability space such that $\sup_{0 \leq t \leq 1} |\overline{U}_n(t) - \overline{W}_o(t)| \xrightarrow{wp1} 0$. O'Reilly shows that

$$\sup_{0 \leq t \leq 1} \left| \frac{\overline{U}_n(t) - \overline{W}_o(t)}{q(t)} \right| \xrightarrow{p} 0 \text{ as } n \to \infty .$$

Thus

$$\sup_{0 \leq t \leq 1} \left| \frac{\overline{U}_n(t)}{q(t)} \right| \xrightarrow{L} \sup_{0 \leq t \leq 1} \left| \frac{W_o(t)}{q(t)} \right| \text{ as } n \to \infty .$$

Moreover, $\sup_{0 \leq t \leq 1} |W_o(t)/q(t)|$ has a (finite) distribution and by the bound given in Lemma 8, (2.21) follows. □

The next result, due to James (1975), characterizes (2.8) for $||\cdot|| = ||\cdot||_{q(F)}$. In accord with James' notation, we let W be the set of positive real-valued functions w on [0,1] such that for some $0 < \delta \leq \frac{1}{2}$, $t^{\frac{1}{2}} w(t)$ is monotone increasing on $(0,\delta]$, $(1-t)^{\frac{1}{2}} w(t)$ is monotone decreasing on $[1-\delta, 1)$, and w is bounded on $[\delta, 1-\delta]$. Here w plays the role of 1/q and values at 0 and 1 are arbitrary.

LEMMA 11. (James). *Let* $U_n(t)$ *be defined as in Lemma* 10. *Let* $w \ \varepsilon \ W$. *If*

(2.22) $$\int_0^1 \frac{w^2(t)}{\log\log\left(\frac{1}{t(1-t)}\right)}\, dt < \infty,$$

*then*

(2.23) $$\overline{\lim_{n\to\infty}} \sup_{0\le t\le 1} \left|\frac{U_n(t)w(t)}{\sqrt{2\log\log n}}\right| = \sup_{0\le t\le 1} [t(1-t)]^{\frac{1}{2}} w(t) \quad w.p.1\,.$$

*Conversely, if* (2.22) *diverges, then the l.h.s. of* (2.23) *is* $+\infty$ *w.p.1.*

Let $Q_4 = \{q\colon\ q=1/w,\ w\in\Omega$ and $w$ satisfies (2.22)$\}$ and $Q_5 = \{q\colon q$ is bounded on $[0,1]$ and $q(t) \ge q^*(t)\ \forall t\in(0,1)$, some $q^*(t) \in Q_4\}$. The following corollary is immediate from Lemmas 8 and 11.

COROLLARY. *Let* $\{X_i\}$ *be a sequence of independent r.v.'s having distribution* F *(not necessarily continuous). Let* $q\in Q_5$. *Then there exists a constant* $M \to \infty$ *such that*

(2.24) $$\overline{\lim_{n\to\infty}} \frac{\sqrt{n}\,||F_n - F||_{q(F)}}{\sqrt{\log\log n}} \le M \quad w.p.1\,.$$

EXAMPLE. For any $0 < \delta_1 < \frac{1}{2}$,

$$q(t) = (t(1-t))^{\frac{1}{2}-\delta_1}$$

belongs to both $Q_2$ and $Q_4$ (and $q^2(t)$ belongs to $Q_1$). Gaenssler and Stute (1976) note that for $q(t) = (t(1-t))^{\frac{1}{2}}$ and $U_n(t)$ defined as in Lemma 10, we have

(2.25) $$\sup_{0<t<1} \left|\frac{U_n(t)}{(t(1-t))^{\frac{1}{2}}}\right| \xrightarrow{P} \infty \quad \text{as } n \to \infty$$

and

(2.26) $$\sup_{0<t<1} \left|\frac{U_n(t)}{\sqrt{2t(1-t)\log\log n}}\right| \xrightarrow{P} 1 \quad \text{as } n \to \infty.$$

However, the last statement of Lemma 11 tells us that the l.h.s. of (2.26) is $+\infty$ w.p.1. Thus a "weak" form of (2.8) holds even though (2.7) and (2.8) do not hold. □

We conclude this section with three theorems which combine the norm theory just discussed with Lemmas 3-5.

THEOREM 1. *Suppose that* T *is continuous at* F *w.r.t.* $||\cdot||_\infty$ *(or w.r.t.* $||\cdot||_{q(F)}$, $q \in Q_1$) *and* $G_F$. *If* (2.5) *and* (2.10) *hold, then*

$$\lim_{n\to\infty} T(F_n) = T(F) \quad w.p.1 . \tag{2.27}$$

THEOREM 2. *Suppose that* T *has a differential at* F *w.r.t.* $||\cdot||_\infty$ *(or w.r.t.* $||\cdot||_{q(F)}$, $q \in Q_3$) *and* $G_F$. *If* (2.5), (2.9), *and* (2.10) *hold, then*

$$\sqrt{n}(T(F_n) - T(F)) \xrightarrow{L} N(0,\sigma^2) \quad as\ n \to \infty . \tag{2.28}$$

THEOREM 3. *Suppose that* T *has a differential at* F *w.r.t.* $||\cdot||_\infty$ *(or w.r.t.* $||\cdot||_{q(F)}$, $q \in Q_5$) *and* $G_F$. *If* (2.5), (2.9), *and* (2.10) *hold, then*

$$\overline{\lim_{n\to\infty}} \frac{\sqrt{n}(T(F_n) - T(F))}{\sqrt{2\sigma^2 \log\log n}} = 1 \quad w.p.1 . \tag{2.29}$$

## 3. FURTHER PRELIMINARIES

Let J be integrable on [0,1]. Then $K(t) = \int_0^t J(u)du$ is absolutely continuous and $K'(t) = J(t)$ wherever J(t) is continuous. For any d.f. F, K(F(x)) is a right continuous function of bounded variation. If J is nonnegative and $\int_0^1 J(u)du = 1$, then K(t) and K(F(x)) are d.f.'s.

The following lemmas will be needed in the proofs of Theorems 4-7. Lemmas 12 and 13 allow the difference, T(F) - T(G), to be written in a convenient form. Lemma 14 establishes a simple inequality. Let $J_1$ be the set of all J integrable on [0,1]. For $J \in \mathcal{J}_1$ let $\mathcal{F}_J = \{F: |\int F^{-1}(t)J(t)dt| < \infty\}$.

LEMMA 12. *If* $F \in \mathcal{F}_J$, *then*

$$\int_0^1 F^{-1}(t)dK(t) = \int_{-\infty}^{\infty} x dK(F(x)) . \tag{3.1}$$

PROOF. Basically we need to justify the substitution $t=F(x)$. Let $I_i = (a_i,b_i]$ be the intervals in [0,1] such that $F^{-1}(a_i,b_i] = x_i$, where $x_i$ is a jump point of $F$ and $F(x_i) = b_i$ and $F(x_i-) = a_i$. Then

$$\int_0^1 F^{-1}(t)dK(t) = \sum_{i=1}^{\infty} \int_{I_i} F^{-1}(t)dK(t) + \int_{[0,1]-\cup I_i} F^{-1}(t)dK(t) . \tag{3.2}$$

Let $S_F$ be the support of $F$ and $A = \{x_1,x_2,\ldots\}$ the set of jump points of $F$. Then

$$\int_{-\infty}^{\infty} x dK(F(x)) = \int_{S_F} x dK(F(x)) = \sum_{i=1}^{\infty} \int_{\{x_i\}} x dK(F(x)) + \int_{S_F - A} x dK(F(x)) . \tag{3.3}$$

We first show that

$$\sum_{i=1}^{\infty} \int_{I_i} F^{-1}(t)dK(t) = \sum_{i=1}^{\infty} \int_{\{x_i\}} x dK(F(x)) . \tag{3.4}$$

Consider the i-th integral of the l.h.s. of (3.4). We have

$$\int_{I_i} F^{-1}(t)dK(t) = x_i \int_{(a_i,b_i]} dK(t) = x_i[K(b_i) - K(a_i)],$$

since $F^{-1}(t) = x_i$ for $t \in (a_i,b_i]$. The i-th integral of the r.h.s. of (3.4) is

$$\int_{\{x_i\}} x dK(F(x)) = x_i[K(F(x_i)) - K(F(x_i-))]$$

$$= x_i[K(b_i) - K(a_i)]$$

by definition of the Stieltjes integral and by substitution of $F(x_i) = b_i$, $F(x_i-) = a_i$. Thus we have term by term equality in (3.4). Now to justify

$$\int_{[0,1]-UI_i} F^{-1}(t)dK(t) = \int_{S_F-A} xdK(F(x)) \tag{3.5}$$

we need only note that $t=F(x)$ is a one-to-one mapping of $S_F$-A onto $[0,1]-UI_i$ and apply a general change of variable lemma (e.g., Dunford and Schwartz, Vol 1, p. 182). Equalities (3.2), (3.3), (3.4), and (3.5) give (3.1). □

Related to the preceding lemma is the familiar identity,

$$\int_0^1 F^{-1}(t)J(t)dt = \int_{-\infty}^{\infty} xJ(F(x))dF(x),$$

which is valid for all continuous F, but not however, for all discontinuous F.

LEMMA 13. *If* $F_1, F_2 \in \mathcal{F}_J$, *then*

$$\int_{-\infty}^{\infty}[K(F_1(x)) - K(F_2(x))]dx = \int_{-\infty}^{\infty} xdK(F_2(x)) - \int_{-\infty}^{\infty} xdK(F_1(x)). \tag{3.6}$$

PROOF. This is a trivial application of integration by parts. The result is well-known when K(t) is a d.f. (e.g., Rao (1973), p. 95 (correcting a misprint)). □

Combining Lemmas 12 and 13, we obtain,

$$T(F) - T(G) = \int[K(G(x)) - K(F(x))]dx \ . \tag{3.7}$$

The next result relates to the quantity $\tilde{V}_{G,F}$ defined for d.f.'s G and F by

$$\tilde{V}_{G,F} = \sup_{G(x)-F(x)\neq 0} \left|\frac{V_{G,F}(x)}{G(x)-F(x)}\right| ,$$

where

$$V_{G,F}(x) = K(G(x)) - K(F(x)) - (G(x) - F(x))J(F(x)) .$$

LEMMA 14. *Suppose that* $J \in \mathcal{J}_1$ *is bounded. Then*

$$\tilde{V}_{G,F} \leq 2||J||_\infty .$$

PROOF.

$$|K(G(x)) - K(F(x))| \leq \left| \int_{F(x)}^{G(x)} J(u)du \right| \leq |G(x) - F(x)| \, ||J||_\infty .$$

Thus, for $G(x) \neq F(x)$,

$$\left| \frac{V_{G,F}(x)}{G(x)-F(x)} \right| \leq \frac{|G(x) - F(x)| \, ||J||_\infty + |G(x) - F(x)| \, |J(x)|}{|G(x)-F(x)|}$$

$$\leq 2||J||_\infty . \square$$

## 4. ROBUST L-FUNCTIONALS

In this section we restrict J to be bounded, continuous a.e. Lebesgue and a.e. $F^{-1}$, and *trimmed* so that $J(u) = 0$ near 0 and 1. Specifically this trimming is of the form

$$J(u) = 0 \qquad u \in [0,t_1) \cup (t_2,1]$$

for $0<t_1<t_2<1$. As noted in Section 1, the functionals generated by these special J functions are generally viewed as *robust* (for justification, see Bickel and Lehmann (1975), I., p. 1054). Moreover, by separating this subclass from the more general functionals of the next section, we will be able to prove the existence of the differential under a minimum of conditions and with respect to the simple sup-norm, $||F_n-F||_\infty = \sup_x |F_n(x)-F(x)|$.

Let F denote a fixed underlying d.f., and define $t_1$ and $t_2$ as above. Note that we may take $F=G_F$ = {all d.f.'s}.

THEOREM 4. *Suppose that*

(4.1) $\qquad J$ *is bounded and continuous a.e. Lebesgue and a.e.* $F^{-1}$;

(4.2) $$J(u) = 0 \text{ for } u \in [0,t_1) \cup (t_2,1].$$

*Then, for* $T(F) = \int F^{-1}(t)J(t)dt$, *the differential of* $T(\cdot)$ *at* $F$ *w.r.t.* $||\cdot||_\infty$ *is given by*

(4.3) $$T(F;\ \Delta) = -\int \Delta(x)J(F(x))dx.$$

REMARKS. (i) Since $T(F;\Delta)$ is clearly linear in $\Delta$, the conclusion may be restated via (2.2) as

(4.4) $$T(G) - T(F) - \int (F(x) - G(x))J(F(x))dx = o(||G-F||_\infty) \text{ as } ||G-F||_\infty \to 0.$$

(ii) If F has bounded support, then (4.2) is not required.

(iii) The "a.e. $F^{-1}$" statement in (4.1) guarantees that J is continuous where F is flat (points F(x) such that x is *not* in the support of F). When J is continuous except at the trimming points $t_1$ and $t_2$, this requirement reduces to the assumption that $t_1$ and $t_2$ correspond to *unique* quantiles of F. □

PROOF OF THEOREM 4. Since $T(F;\Delta)$ is linear, we need only show (4.4). By (3.7) the l.h.s. of (4.4) can be written as

$$-\int [K(G(x)) - K(F(x)) - (G(x)-F(x))J(F(x))]dx = -\int V_{G,F}(x)dx\ .$$

Let (a,b) be such that $G(a) < t_1$, $F(a) < t_1$, $G(b) > t_2$, $F(b) > t_2$. Then, since $J(F(x))$ and $K(G(x)) - K(F(x))$ are 0 outside (a,b),

$$\int V_{G,F}(x)dx = \int_a^b V_{G,F}(x)\, dx .$$

Let B = {x: F(x) is a discontinuity point of J} and define

$$W_{G,F}(x) = \frac{V_{G,F}(x)}{G(x)-F(x)} \qquad \text{if } G(x) \neq F(x)$$

$$= 0 \qquad \text{if } G(x) = F(x) .$$

Since B is a Lebesgue-null set (using the fact that J is continuous a.e. $F^{-1}$), we have

$$\left|\int V_{G,F}(x)dx\right| = \left|\int_{(a,b)-B} V_{G,F}(x)dx\right|$$

$$= \left|\int_{(a,b)-B} (G(x) - F(x))(W_{G,F}(x))dx\right|$$

$$\leq ||G-F||_\infty \int_{(a,b)-B} |W_{G,F}(x)|dx.$$

By definition of the derivative,

$$\lim_{||G-F||_\infty \to 0} \left|W_{G,F}(x)\right| = 0 \qquad \forall\ x \in (a,b) - B$$

because $K'(F(x) = J(F(x))$ on $(a,b) - B$. Since $\left|W_{G,F}(x)\right| \leq \tilde{V}_{G,F} \leq 2||J||_\infty$ by Lemma 14, we can justify interchange of the operations of limit and integration through use of the theorem on bounded convergence for a finite interval. That is,

$$\lim_{||G-F||_\infty \to 0} \frac{|\int V_{G,F}(x)dx|}{||G-F||_\infty} \leq \lim_{||G-F||_\infty \to 0} \int_{(a,b)-B} \left|W_{G,F}(x)\right|dx$$

$$= \int_{(a,b)-B} \lim_{||G-F||_\infty \to 0} \left|W_{G,F}(x)\right|dx = 0. \quad \square$$

Asymptotic normality and an LIL follow easily from Theorems 2, 3, and 4. Note that for the differential defined in Theorem 4, we have

(4.5) $$T[F;x] = T(F;\delta_x - F) = \int [F(t) - I(x \le t)]J(F(t))dt.$$

Thus

(4.6) $$E_F\{T[F;X]\} = 0 ,$$

and

(4.7) $$\sigma^2 = \mathrm{Var}_F\{T[F;X]\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J(F(s))J(F(t))(F(\min(s,t)) - F(s)F(t))dsdt.$$

COROLLARY. *Suppose that* J *and* F *satisfy* (4.1) *and* (4.2) *and that* $\sigma^2 > 0$. *Let* $\{X_i\}$ *be a sequence of independent r.v.'s having distribution* F. *Then*

(4.8) $$\sqrt{n}(T(F_n) - T(F)) \xrightarrow{L} N(0,\sigma^2) \quad as\ n \to \infty ,$$

*and*

(4.9) $$\overline{\lim_{n\to\infty}} \frac{\sqrt{n}(T(F_n) - T(F))}{\sqrt{2\sigma^2 \log\log n}} = 1 \quad w.p.1 .$$

PROOF. Condition (2.9) is given by (4.6) and the assumption that $\sigma^2 > 0$. Condition (2.5) is satisfied since $G_F$ in this case is the set of all d.f.'s. Condition (2.10) holds by hypothesis and Theorem 4 provides the existence of the differential. Thus the conditions of Theorems 2 and 3 are satisfied. □

We could formulate the above corollary for dependent variables also. Specifically, for (4.8), we could replace the independence assumption by the assumption that the $\{X_i\}$ satisfy (2.11) and (2.7) with $||\cdot|| = ||\cdot||_\infty$. For (4.9), we would need (2.12) and (2.8) with $||\cdot|| = ||\cdot||_\infty$.

Note that (4.9) yields strong consistency of $T(F_n)$. However, by appeal to continuity of T, we can slightly relax (4.1).

THEOREM 5. *Suppose that* J *is integrable on* [0,1] *and bounded and that* (4.2) *holds. Then* T *is continuous at* F *w.r.t.* $||\cdot||_\infty$. *Further, if* $\{X_i\}$ *is a sequence of independent r.v.'s having distribution* F, *then*

$$\lim_{n\to\infty} T(F_n) = T(F) \quad w.p.1 \tag{4.10}$$

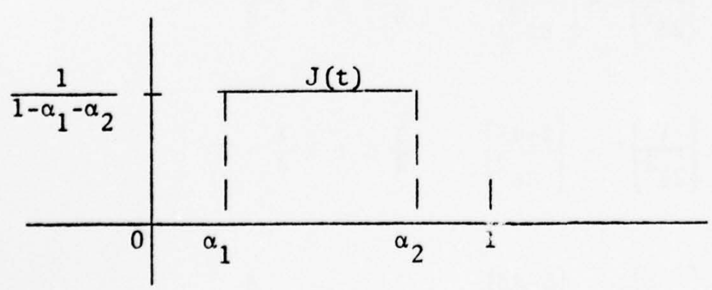
PROOF. Let (a,b) be as in the proof of Theorem 4. Then

$$\begin{aligned}|T(G)-T(F)| &= \left|\int_a^b [K(G(x)) - K(F(x))]dx\right| \\ &\le |b-a|\ ||J||_\infty\ ||G-F||_\infty .\end{aligned}$$

Thus T is continuous at F w.r.t. $||\cdot||_\infty$, and an appeal to Theorem 1 yields (4.10). □

EXAMPLES. (i) The *trimmed mean.*

$$\begin{aligned}J(t) &= \frac{1}{1-\alpha_1-\alpha_2} && \alpha_1 \le t \le \alpha_2 \\ &= 0 && \text{o.w.}\end{aligned}$$



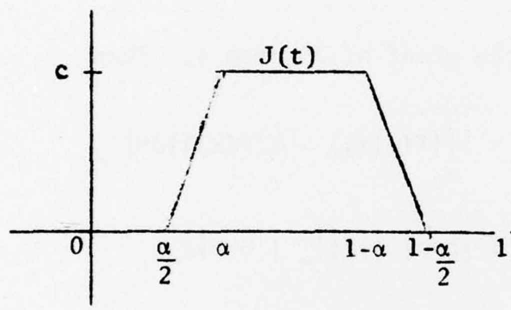

(ii) The *smoothly trimmed mean* (from Stigler (1973)). Let $\alpha_1 = 1-\alpha_2 = \alpha$ and c be a constant.

$$J(t) = \left(t-\frac{\alpha}{2}\right)\left(\frac{2c}{\alpha}\right), \quad \frac{\alpha}{2} \le t \le \alpha$$

$$= c, \quad \alpha \le t \le 1-\alpha$$

$$= \left(1-\frac{\alpha}{2}-t\right)\left(\frac{2c}{\alpha}\right), \quad 1-\alpha \le t \le 1-\frac{\alpha}{2}$$

$$= 0, \quad \text{o.w.}$$




(iii) Consider the *interquartile range*, $\frac{F^{-1}(3/4)-F^{-1}(1/4)}{2}$. This functional is obtained from (1.1) by letting $K(t)$ place mass $\frac{1}{2}$ at $F^{-1}(3/4)$ and $-\frac{1}{2}$ at $F^{-1}(1/4)$. Although Theorem 4 doesn't apply directly to discrete $K(t)$ we can approximate $\frac{F^{-1}(3/4)-F^{-1}(1/4)}{2}$ by a functional with absolutely continuous $K(t)$. Let

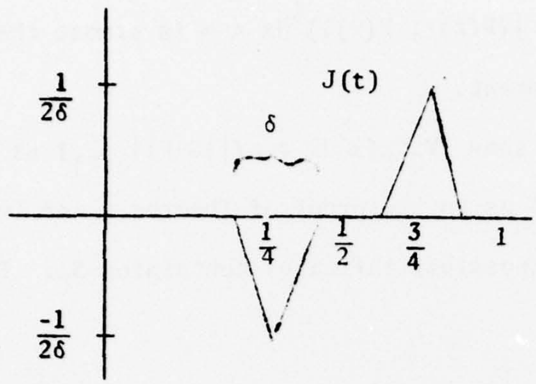
$$J(t) = \left(\frac{-1}{2\delta^2}\right)t + \left(\frac{1-4\delta}{8\delta^2}\right) \qquad \frac{1}{4}-\delta \le t \le \frac{1}{4}$$

$$= \left(\frac{1}{2\delta^2}\right)t - \left(\frac{1+4\delta}{8\delta^2}\right) \qquad \frac{1}{4} \le t \le \frac{1}{4}+\delta$$

$$= \left(\frac{1}{2\delta^2}\right)t - \left(\frac{3-4\delta}{8\delta^2}\right) \qquad \frac{3}{4}-\delta \le t \le \frac{3}{4}$$

$$= \left(\frac{-1}{2\delta^2}\right)t + \left(\frac{3+4\delta}{8\delta^2}\right) \qquad \frac{3}{4} \le t \le \frac{3}{4}+\delta$$

$$= 0 \qquad \text{o.w.}$$




The area within each spike is $\frac{1}{2}$. Note that for this functional as well as for the smoothly trimmed mean, J is continuous everywhere. Thus Theorem 4 applies to these functionals for *any* d.f. F.

## 5. GENERAL L-FUNCTIONALS

In this section we remove the trimming restriction on J. In order to deal with the weight placed on the extremes of F, we will use the q-norms introduced in Section 2. Let

$$\mathcal{F} = \{F: \ |\int F^{-1}(t)J(t)dt| < \infty\} \text{ and } \mathcal{G}_F = \{G: \ G \in \mathcal{F} \text{ and } S_G \subset S_F\},$$

where $S_F$ is the support of F. Let q be a bounded positive function on (0,1).

THEOREM 6. *Suppose that* J, q, *and* $F \in \mathcal{F}$ *satisfy* (4.1) *and*

$$\int q(F(x))dx < \infty \ . \tag{5.1}$$

*Then the differential of* $T(F) = \int F^{-1}(t)J(t)dt$ *at* F *w.r.t.* $||\cdot||_{q(F)}$ *and* $\mathcal{G}_F$ *is given by* (4.3).

REMARKS. Condition (5.1) governs the tails of F. For $q(t) = [t(1-t)]^{\frac{1}{2}-\delta_1}$, (5.1) becomes

$$\int (F(x)(1-F(x)))^{\frac{1}{2}-\delta_1}dx < \infty \ . \tag{5.2}$$

Stigler (1974, p. 686) notes that $\int (F(x)(1-F(x)))^{\frac{1}{2}}dx < \infty$ is almost the same as the existence of a finite 2nd movement. □

PROOF OF THEOREM 6. We must show $\int V_{G,F}(x)dx = o(||G-F||_{q(F)})$ as $||G-F||_{q(F)} \to 0$, $G \in \mathcal{G}_F$. Define B as in the proof of Theorem 4 and let the closure of $(x_1,x_2)$ be the smallest interval (possibly infinite) containing $S_F$. Then for $G \in \mathcal{G}_F$ we have

$$\left|\int V_{G,F}(x)dx\right| = \left|\int_{(x_1,x_2)-B} V_{G,F}(x)dx\right|$$

$$= \left|\int_{(x_1,x_2)-B} \left(\frac{G(x)-F(x)}{q(F(x))}\right)\left(W_{G,F}(x)\right) q(F(x))dx\right|$$

$$\leq ||G-F||_{q(F)} \int_{(x_1,x_2)-B} \left|W_{G,F}(x)\right| q(F(x))dx.$$

Once again we have only to show that the interchange of limit and integration is valid, since

$$\lim_{\substack{||G-F||_{q(F)}\to 0 \\ G\in\mathcal{G}_F}} |W_{G,F}(x)| = 0 \qquad \forall\, x \in (x_1,x_2) - B.$$

(Recall that for $G \in \mathcal{G}_F$, $|G(x)-F(x)| \leq ||G-F||_\infty \leq ||q||_\infty ||G-F||_{q(F)} \forall x$.) In this case we appeal to dominated convergence by way of Lemma 14. That is, by Lemma 14,

$$|W_{G,F}(x)|q(F(x)) \leq \tilde{V}_{G,F}q(F(x)) \leq 2||J||_\infty\, q(F(x)),$$

and the r.h.s. is integrable by (5.1). Thus for $G \in \mathcal{G}_F$, we have

$$\lim_{||G-F||_{q(F)}\to 0} \frac{|\int V_{G,F}(x)dx|}{||G-F||_{q(F)}} \le \lim_{||G-F||_{q(F)}\to 0} \int_{(x_1,x_2)-B} |W_{G,F}(x)|q(F(x))dx$$

$$= \int_{(x_1,x_2)-B} \lim_{||G-F||_{q(F)}\to 0} |W_{G,F}(x)|q(F(x))dx = 0. \square$$

Note that unbounded J's could be allowed if the above interchange of operations could be justified in such a situation. Asymptotic normality and an LIL are provided by the following corollary.

COROLLARY. *Suppose that* $\sigma^2$ *defined by* (4.7) *is finite and positive. Let* $\{X_i\}$ *be a sequence of independent r.v.'s having distribution* F. *If* $q \in Q_3$ *and* J *and* $F \in \mathcal{F}$ *satisfy* (4.1) *and* (5.1), *then* (4.8) *holds. If* $q \in Q_5$ *and* J *and* $F \in \mathcal{F}$ *satisfy* (4.1) *and* (5.1), *then* (4.9) *holds.*

The following analogue of Theorem 5 gives continuity of T and strong consistency of $T(F_n)$.

THEOREM 7. Let $q \in Q_1$. *Suppose that* J *is integrable on* [0,1] *and bounded and that* (5.1) *holds. Then* T *is continuous at* F *w.r.t.* $||\cdot||_{q(F)}$. *Further, if* $\{X_i\}$ *is a sequence of independent r.v.'s having distribution* F, *then* (4.10) *holds.*

PROOF.

$$|T(G)-T(F)| = |\int[K(G(x))-K(F(x))]dx|$$

$$\le ||J||_\infty \int|G(x)-F(x)|dx$$

$$\le ||J||_\infty ||G-F||_{q(F)} \int q(F(x))dx. \square$$

EXAMPLES. (i) The *mean*, $J(t)\equiv 1$. In this case Lemma 12 yields the familiar identity

$$\int_0^1 F^{-1}(t)dt = \int_{-\infty}^{\infty} xdF(x).$$

(ii) Gini's *mean difference*, $J(t) = t - \frac{1}{2}$.

(iii) A location estimator suggested by Bickel (1973), $J(t) = 6t(1-t)$.

## 6. EXTENSION TO A LARGER CLASS OF L-FUNCTIONALS

Some authors consider the slightly more general statistic $\sum_{i=1}^{n} c_{in} h(X_{in})$ (e.g. Chernoff, Gastwirth, and Johns(1967)). Similarly, our functional is

$$(6.1) \qquad T(F) = \int_0^1 h(F^{-1}(t))dK(t) .$$

Since $h(x) = x$ is commonly the function used in applications, attention has been confined to this case in Sections 3-5. However, it is easy to extend the results of those sections to T defined by (6.1). Let $H_1$ be the set of continuous functions h defined on $(-\infty,\infty)$ such that $h = h^+ - h^-$ for monotone increasing functions $h^+$ and $h^-$. For $h \in H_1$ let $F_h = \{F: \ |\int h(F^{-1}(t))dK(t)| < \infty\}$. First we give without proof the analogues to Lemmas 12 and 13.

LEMMA 12*. *If* $F \in F_h$, *then*

$$\int_0^1 h(F^{-1}(t))dK(t) = \int_{-\infty}^{\infty} h(x)dK(F(x)).$$

LEMMA 13*. *If* $F_1, F_2 \in F_h$, *then*

$$\int (K(F_1(x))-K(F_2(x)))dh(x) = \int h(x)dK(F_2(x)) - \int h(x)dK(F_1(x)).$$

Let $\mu_h$ be the measure corresponding to $h^+ + h^-$ and let B be defined as in the proof of Theorem 4. The following are analogues to Theorems 4 and 6. We omit the proofs as well as the analogues to Theorems 5 and 7.

THEOREM 4.* *Suppose that* J, h, *and* $F \in F_h$ *satisfy* (4.2) *and*

$$(6.2) \qquad J \textit{ is bounded and continuous a.e. } \mu_h \textit{ and } \mu_h(B) = 0.$$

*Then* T *defined by* (6.1) *has a differential at* F *w.r.t.* $||\cdot||_\infty$ *given by*

$$T(F;\Delta) = -\int \Delta(x) J(F(x)) dh(x). \tag{6.3}$$

THEOREM 6.* *Suppose that* J, h, q, *and* $F \in \mathcal{F}_h$ *satisfy* (6.2) *and*

$$\left|\int q(F(x)) dh(x)\right| < \infty. \tag{6.4}$$

*Then* T *defined by* (6.1) *has a differential at* F *w.r.t.* $||\cdot||_{q(F)}$ *and* $\mathcal{G}_F = \{G\colon G \in \mathcal{F}_h$ *and* $S_G \in S_F\}$ *given by* (6.3).

Bickel and Lehmann (1976, III) discuss the functional

$$\tau(F) = \left[\int_0^1 (F_\mu^{-1}(t))^\alpha dK(t)\right]^{\frac{1}{\alpha}}, \tag{6.5}$$

where $F_\mu$ denotes the distribution of $|X-\mu|$ when X has distribution F, $\mu$ is a *known* constant, $\alpha > 0$, and K(t) is a d.f. on (0,1). As a functional of F, (6.5) is not amenable to our methods. However, if we let $\tau(F) = T(F_\mu)$, a functional of $F_\mu$, it is easy to see that $[T(F_\mu)]^\alpha$ has the appropriate form, with $h(x) = x^\alpha$. Then, assuming the conditions of Theorem 4* (or Theorem 6*), we have that the differential of $[T(\cdot)]^\alpha$ at $F_\mu$ w.r.t. $||\cdot||_\infty$ (or w.r.t. $||\cdot||_{q(F)}$ and $\mathcal{G}_{F_\mu}$) is given by

$$T(F_\mu;\Delta) = -\int \Delta(x) J(F_\mu(x)) d(x^\alpha) . \tag{6.6}$$

Since we really want the differential of $T(\cdot)$ rather than the differential of $[T(\cdot)]^\alpha$, the following "chain rule" is required.

LEMMA 15. *Let* f(x) *be a real-valued function,* f: R→R, *with a derivative at* x=T(F), f'(T(F)). *Suppose further that the functional* $T(\cdot)$ *has a differential at* F *w.r.t.* $||\cdot||$ *and* $\mathcal{G}_F$ *given by* T(F; G-F), *and* that T(G) - T(F) *is*

$O(||G-F||)$ *as* $||G-F||\to 0$. *Then the composite functional* $S(\cdot) = f(T(\cdot))$ *has a differential w.r.t.* $||\cdot||$ *and* $G_F$ *given by*

$$S(F;\Delta) = f'(T(F))T(F;\Delta). \tag{6.7}$$

PROOF. Trivial.

We apply this lemma with $f(x) = x^{1/\alpha}$, so that $f'(x) = (x^{1/\alpha-1})/\alpha$. Combining (6.6) and (6.7), we obtain

$$T(F_\mu;G-F_\mu) = \frac{[T(F_\mu)]^{1-\alpha}}{\alpha} \cdot \int_0^\infty (F_\mu(x)-G(x))J(F_\mu(x))d(x^\alpha).$$

EXAMPLE. Let $(\mu,\sigma^2)$ be the mean and variance of $F$, $\alpha=2$, and $K(t)=t$. Then $T(F_\mu)$ is the standard deviation $\sigma$ of $F$, and the differential is

$$T(F_\mu;G-F_\mu) = \frac{\int_0^\infty (F_\mu(x)-G(x))d(x^2)}{2T(F_\mu)} = \frac{\int_0^\infty x^2 d(G(x)-F_\mu(x))}{2T(F_\mu)}. \tag{6.8}$$

For a sample $X_1,\ldots,X_n$ of independent r.v.'s with distribution $F$, let $Y_i=|X_i-\mu|$ and $F_{\mu n}$ be the sample d.f. formed from the $Y_i$'s. Then substitution of $G=F_{\mu n}$ in (6.8) yields

$$T(F_\mu;F_{\mu n}-F_\mu) = \frac{\int x^2 d(F_{\mu n}(x)-F_\mu(x))}{2\sigma} = \frac{1}{n}\sum_{i=1}^n\left(\frac{Y_i^{\,2}-\sigma^2}{2\sigma}\right).$$

Since $E(Y_i^{\,2}-\sigma^2) = 0$, and $E(Y_i^{\,2}-\sigma^2)^2 = E(X-\mu)^4 - \sigma^4$, the central limit theorem yields (via Lemma 2) that

$$\sqrt{n}(T(F_{\mu n}) - \sigma) \xrightarrow{L} N\left(0, \frac{E(x-\mu)^4-\sigma^4}{4\sigma^2}\right) \text{as } n\to\infty.$$

This last result could have been anticipated since $T(F_{\mu n})$ turns out to be the usual estimator of $\sigma$:

$$T(F_{\mu n}) = [\int_0^\infty x^2 dF_{\mu n}(x)]^{\frac{1}{2}} = [\frac{1}{n}\sum_{i=1}^{n}\int_0^\infty x^2 d(I(-x\leq X_i-\mu\leq x))]^{\frac{1}{2}}$$

$$= [\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2]^{\frac{1}{2}}.$$

## 7. CONCLUSIONS AND COMPARISONS

Stigler (1974) provides good motivation for the use of L-statistics which are generated by smooth weight functions J. One inherent value of these statistics is that the theorems justifying their use (asymptotic normality, etc.) place the force of the restricting conditions on J (over which we have control) rather than on F, which is often only partially known. Thus, practicing statisticians can actually *verify*, rather than assume, most of the needed hypotheses. From consideration of counterexamples, it appears that J continuous a.e. $F^{-1}$ is essentially a necessary condition for asymptotic normality (c.f. Stigler (1974), Section 5.6). Requiring J to be bounded is not very restrictive, since in most situations, unbounded J functions would produce notoriously nonrobust estimators. Hence, J bounded and continuous a.e. $F^{-1}$ is a natural restriction which we and Stigler (1974) have in common. The additional requirement that J be continuous a.e. Lebesgue is necessary for our Theorems 4 and 6 and Stigler needs it for several of his theorems (c.f. Stigler (1974), Theorems 3 and 4). It should be noted that Stigler's statistic, defined by

$$S_n = \frac{1}{n}\sum_{i=1}^{n} J\left(\frac{i}{n+1}\right)X_{in} ,$$

is somewhat different from our (1.3). Nevertheless, both converge to $\int F^{-1}(t)J(t)dt$, and rough comparisons can be made. Basically, Stigler's Theorems 2 and 5, which establish

$$(7.1) \qquad \frac{S_n - E(S_n)}{\sigma(S_n)} \xrightarrow{L} N(0,1) \text{ as } n \to \infty ,$$

require a little less than our Theorems 4 and 6. However, for practical use, (7.1) is needed with $E(S_n)$ replaced by $\int F^{-1}(t)J(t)dt$ and $\sigma(S_n)$ replaced by $\sigma$ defined by (4.6). To accomplish this, Stigler requires stronger assumptions on J than our (4.1) (see Theorem 4, Stigler (1974)). On the other hand, the tail condition (5.2) is not quite as mild as $\int[(F(x))(1-F(x))]^{\frac{1}{2}}dx < \infty$, the one given in Stigler's Theorem 4. Stigler extends to certain independent but non-identically distributed variables, whereas we can extend to identically distributed but dependent variables. Of course, our method also establishes strong consistency and an LIL without additional assumptions.

The results of Gregory (1976), like ours, are obtained by differential methods. He proves Theorem 4* for J bounded and continuous and F absolutely continuous. It appears that his use of the chain rules associated with formal Frechet differentiation requires J to be continuous everywhere. This indicates the power of our more flexible version of the differential approach. Weakening of the absolutely continuous assumption on F could be made via our Lemma 8. Also, his special q function yields a little sharper asymptotic normality result, though it doesn't appear to belong to $Q_3$ or $Q_5$.

Wellner (1977a), (1977b) provides somewhat more general results for strong consistency and the LIL. In particular, his LIL is of the Strassen type and allows combinations of quantiles as well.

We see that Theorems 4-7 provide improved results with regard to asymptotic normality of L-statistics and additional results for strong consistency and the LIL. Furthermore, applications to different situations involving dependence are straightforward as the relevant theory concerning $||F_n-F||$ becomes available.


## ACKNOWLEDGEMENT

I am indebted to Professor R. J. Serfling for suggesting the problem and for careful reading of the manuscript as well as providing many helpful comments. I would also like to thank Professor Gavin Gregory for providing a copy of his manuscript and Professor George Policello for detecting an error in an earlier proof of Theorem 6.

## REFERENCES


[1] Bickel, P. J. (1967). Some contributions to the theory of order statistics. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* 1 575-591. Univ. of California Press.

[2] Bickel, P. J. (1973). On some analogues of linear combinations of order statistics in the linear model. *Ann. Statist.* 1 597-616.

[3] Bickel, P. J. and Lehmann, E. L. (1975). Descriptive statistics for nonparametric models. I. Introduction. *Ann. Statist.* 3 1038-1044; II. Location. *Ann. Statist.* 3 1045-1069.

[4] Bickel, P. J. and Lehmann, E. L. (1976). Descriptive statistics for nonparametric models. III. Dispersion. *Ann. Statist.* 4 1139-1158.

[5] Boos, D. D. (1977). The functional approach in statistical theory and robust inference. Unpublished dissertation, Florida State University.

[6] Boos, D. D. and Serfling, R. J. (1977) On differentials, asymptotic normality and almost sure behavior of statistical functions, with application to M-statistics for location. FSU Statistics Report 415 (ONR Technical Report No. 115), Florida State University, Tallahassee.

[7] Chernoff, H., Gastwirth, J. L. and Johns, M. V. (1967). Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. *Ann. Math. Statist.* 38 52-72.

[8] Dunford, N. and Schwartz, J. T. (1958). *Linear Operators*. Vol. I. Interscience Publishers, Inc., New York.

[9] Gaenssler, P. and Stute, W. (1976). A survey on some results for empirical processes in the i.i.d. case. University of Bochum.

[10] Gregory, G. G. (1976). Large sample theory of differantiable statistical functions. Unpublished manuscript.

[11] Hampel, F. R. (1971). A general qualitative definition of robustness. *Ann. Math. Statist.* 42 1887-1896.

[12] Hampel, F. R. (1974). The influence curve and its role in robust estimation. *J. Amer. Statist. Assoc.* 69 383-393.

[13] Huber, P. J. (1972). The 1972 Wald Lecture. Robust statistics: a review. *Ann. Math. Statist.* 43 1041-1067.

[14] James, B. R. (1975). A functional law of the iterated logarithm for weighted empirical distributions. *Ann. Prob.* 3 762-772.

[15] O'Reilly, N. E. (1974). On the weak convergence of empirical processes in sup-norm metrics. *Ann. Prob.* 4 642-651.

[16] Pyke, R. and Shorack, G. R. (1968). Weak convergence of a two-sample empirical process and a new approach to Chernoff-Savage Theorems. *Ann. Math. Statist.* 39 755-771.

[17] Rao, C. R. (1973). *Linear Statistical Inference and its Applications.* Wiley, New York.

[18] Shorack, G. R. (1969). Asymptotic normality of linear combinations of functions of order statistics. *Ann. Math. Statist.* 40 2041-2050.

[19] Shorack, G. R. (1972). Functions of order statistics. *Ann. Math. Statist.* 43 412-427.

[20] Stigler, S. M. (1969). Linear functions of order statistics. *Ann. Math. Statist.* 40 770-788.

[21] Stigler, S. M. (1973). The asymptotic distribution of the trimmed mean. *Ann. Statist.* 1 472-477.

[22] Stigler, S. M. (1974). Linear functions of order statistics with smooth weight functions. *Ann. Statist.* 2 676-693.

[23] Wellner, J. A. (1977a). A Glivenko-Cantelli theorem and strong laws of large numbers for functions of order statistics. *Ann. Statist.* 5 473-480.

[24] Wellner, J. A. (1977b). A law of the iterated logarithm for functions of order statistics. *Ann. Statist.* 5 481-494.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

| | | |
|---|---|---|
| 1. REPORT NUMBER<br><br>ONR Technical<br>Report No. 116 | 2. GOVT. ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and subtitle)<br><br>A Differential for L-statistics | | 5. TYPE OF REPORT & PERIOD COVERED<br><br>Technical Report<br><br>6. PERFORMING ORG. REPORT NUMBER<br><br>FSU Statistics Report M416 |
| 7. AUTHOR(s)<br><br>Dennis D. Boos | | 8. CONTRACT OR GRANT NUMBER<br><br>Office of Naval Research<br>Contract No. N00014-76-C-0608 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br><br>Florida State University<br>Department of Statistics<br>Tallahassee, Florida 32306 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br><br>Office of Naval Research<br>Statistics & Probability Program<br>Arlington, Virginia 22217 | | 12. REPORT DATE<br><br>June, 1977<br><br>13. NUMBER OF PAGES<br><br>37 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) | | 15. SECURITY CLASS (of this report)<br><br>Unclassified<br><br>15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from report).

18. SUPPLEMENTARY NOTES

19. KEY WORDS

L-Statistics; Differentials; Asymptotic Normality; Almost Sure Behavior.

REPORT DOCUMENTATION PAGE

20. ABSTRACT


A DIFFERENTIAL FOR L-STATISTICS

Let $X_{1n} \leq X_{2n} \leq \ldots \leq X_{nn}$ be an ordered sample from a distribution F and $c_{in}$ a sequence of constants. Statistics of the form $\sum_{i=1}^{n} c_{in} X_{in}$ are called "linear functions of order statistics," "L-estimators," or simply "L-statistics." Various methods of generating the constants $c_{in}$ have been considered, including $\tilde{c}_{in} = n^{-1}J(i/n+1)$ or $n^{-1}J(1/n)$ and $\hat{c}_{in} = \int_{(i-1)/n}^{i/n} J(u)du$ for fixed "score" functions J. L-statistics of the form $T_n = \sum_{i=1}^{n} \hat{c}_{in} X_{in}$ can be obtained from the functional $T(F) = \int F^{-1}(t)J(t)dt$ by substitution of the sample d.f. $F_n$ for F. Under the mild assumption that J is bounded and continuous a.e. Lebesgue and a.e. $F^{-1}$ and under a tail restriction on F of the form $\int q(F(x))dx < \infty$ (e.g., $q(t) = [t(1-t)]^{\frac{1}{2}-\delta}$, $0 < \delta < \frac{1}{2}$), it is shown that $T(\cdot)$ has a Frechet-type differential. The tail restriction may be dropped if J trims the extremes. In either case it follows that if $\{X_i\}$ is a sequence of independent observations on F, then $\sqrt{n}(T(F_n)-T(F))$ is asymptotically normal and obeys a law of the iterated logarithm. Continuity of T holds under milder conditions on J and F. This leads to strong consistency, $T(F_n) \xrightarrow{wpl} T(F)$. No continuity restrictions are imposed on F, so that the results are applicable to a wide class of distributions of interest in robust estimation. Illustration is provided by examples including the trimmed mean, the smoothly trimmed mean, and approximations to the interquartile range. The asymptotic normality result is competitive with one of Stigler (1974) for the closely related statistic $S_n = \sum_{i=1}^{n} \tilde{c}_{in} X_{in}$, obtained under stronger conditions on J but a slightly milder condition on F. However, in addition to asymptotic normality of $T(F_n)$, the differential approach of the present paper yields characterization of the almost sure behavior of $T(F_n)$ and lends itself to straightforward extension to the case of dependent variables.


F sub n

S sub n = Sum from i=(1 to n) of (c wiggle sub in)(X sub in)